Tridiagonal random matrix: Gaussian fluctuations and deviations

Deng Zhang[1 2]

Abstract. This paper is devoted to the Gaussian fluctuations and deviations of the traces of tridiagonal random matrix. Under quite general assumptions, we prove that the traces are approximately normal distributed. Multi-dimensional central limit theorem is also obtained here. These results have several applications to various physical models and random matrix models, such as the Anderson model, the random birth-death Markov kernel, the random birth-death Q matrix and the β-Hermite ensemble. Furthermore, under some independent and identically distributed condition, we also prove the large deviation principle as well as the moderate deviation principle for the traces.

Keyword: Central limit theorem, moderate deviation, large deviation, tridiagonal random matrix.

2010 Mathematics Subject Classification: 60B20, 60F05, 60F10.

[1]Department of Mathematics, Shanghai Jiao Tong University, 200240 Shanghai, China. This work is supported by Notional Basic Research Program of China (973 Program, 2011CB808000, 2015CB856004) and Notional Natural Science Fundations of China (11171216).

[2]Email address: zhangdeng@amss.ac.cn

1 Introduction

We are here concerned with the tridiagonal random matrix

$$Q_n = \begin{pmatrix} d_1 & b_1 & & & & \\ a_1 & d_2 & b_2 & & & \\ & a_2 & d_3 & b_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-2} & d_{n-1} & b_{n-1} \\ & & & & a_{n-1} & d_n \end{pmatrix}, \tag{1.1}$$

where a_i, b_i and d_i, $1 \le i \le n$, are random variables with $a_0 = b_n = 0$.

Tridiagonal random matrix attracts significant interests in various fields. In quantum mechanics, it is a finite-difference model of one-dimensional random Schrödinger operator, such as the extensively studied Anderson model where $a_i = b_i = -1$ and all d_i are independent and identically distributed (see e.g. [28, 7]). The non-symmetric model with $a_i/b_i > 0$ also arises in the non-Hermitian quantum mechanics of Hatano and Nelson, see e.g. [21] and references therein. Tridiagonal random matrix is also a basic model of random walks with random environment in chain graphs, interesting examples include the random birth-death Markov kernel proposed in [5, 8], where $a_{i-1} + d_i + b_i = 1$ and $\{(a_{i-1}, d_i, b_i)\}$ is an ergodic random field, and the random birth-death Q matrix recently studied in [23], where $d_i = -(a_{i-1} + b_i)$ and $\{a_i\}$ and $\{b_i\}$ are two sequences of strictly stationary ergodic processes.

Moreover, tridiagonal random matrix also plays an important role in random matrix theory. One well-known model is the β-Hermite ensemble, in which $a_i = b_i$, a_i is distributed as $\beta^{-1/2}\chi_{i\beta}$ ($\chi_{i\beta}$ is the χ distribution with $i\beta$ degrees of freedom), $\beta > 0$, d_i is normally distributed as $N(0, 2/\beta)$, and $\{a_i, d_i\}$ are independent. This model was first proposed in [17] and generalized the classical Gaussian ensembles for $\beta = 1, 2, 4$, corresponding to the Gaussian orthogonal ensemble (GOE), Gaussian unitary ensemble (GUE) and Gaussian symplectic ensemble (GSE) respectively. Due to its simple tridiagonal structure, one expects to investigate some interesting spectral phenomena of general random matrices from the study of the tridiagonal random matrix (1.1). For the generalization of the β-Hermite ensemble to general symmetric tridiagonal random matrix (with independent entries), we refer the interested reader to [29], which is another motivation for the present work.

The interest of this paper lies in the Gaussian fluctuations and deviations of the trace of powers, or in other words, the moments of empirical spectral distribution of the tridiagonal random matrix (1.1).

The fluctuations of the traces of random matrices are extensively studied in the literature and, in a general context, turn out to be Gaussian. For instance, the traces of the classical (unitary, orthogonal and symplectic) compact groups were proved by Diaconis and Shahshahani [14], by using the representation theory, to be independent and normally distributed in the limit. For the Wigner matrices, the Gaussian fluctuations of the traces are presented in [1, Theorem 2.1.31] (see also [22, Theorem 2.1]), based on elaborate computations of moments. See also [29] for the symmetric tridiagonal random matrix with indepenent entries, where the Gaussian fluctuations of the traces were obtained by a judicious counting of levels of paths. We also refer to [2, 18, 26, 25, 30, 31] and references therein for fluctuations of other linear eigenvalue statistics.

Furthermore, large deviation was well-known for the empirical spectral distribution of the unitary invariant ensembles, including the β-Hermite ensemble (see [3] and [1]). See also [4] for the non self-ajoint matrix with independent Gaussian entries. Moderate deviation principle for the empirical spectral distribution of the Gaussian divisible matrix was studied in [12]. Moreover, for the eigenvalue counting function of the determinantal point process and Wigner matrix, we refer to [15] and [16] .

Motivated by the physical models and random matrix models mentioned above, we here consider the tridiagonal random matrix (1.1) under quite general assumptions (see $(H.1)$ and $(H.2)$ in Section 3, see also $(H.3)$ in Section 4), which, in particular, allow the non-independent entries in (1.1). We prove the Gaussian fluctuations of the traces of such model. Multi-dimensional central limit theorem is also given. Moreover, the large deviations of the traces, in relation to those of additive functionals of uniform Markov chains, as well as the moderate deviation results are also obtained here. These results are applicable to various models, such as the Anderson model, the random birth-death Markov kernel, the random birth-death Q matrix as well as the symmetric tridiagonal random matrix, including the β-Hermite ensemble.

Our proof is quite different from the standard method of moments mentioned above and relies crucially on a path expansion of the trace (see (2.5) below), which is formulated according to the types of circuits determined by the tridiagonal structure of (1.1). This formula was recently used in our arti-

cle [23] to study the limiting spectral distribution of the random birth-death Q matrix. The advantage of this formula is that, it reveals that the fluctuations and deviations of the traces are asymptotically the same to those of the sum of m-dependent random variables. This new point of view gives a unified way to understand the limit behavior of traces of quite general tridiagonal random matrices with non-symmetric structure or non-independent entries. Moreover, it also allows to employ the analytic tools (e.g. blocking arguments, Lyapunov's central limit theorem and large deviations of additive functionals of Markov chains) to obtain the Gaussian fluctuations and deviations of the traces.

The remainder of this paper is organized as follows. In Section 2 we first set up some preliminary notations and definitions, then we present the path expansion formula (2.5). Section 3 is devoted to the Gaussian fluctuations of the traces, and Section 4 is concerned with the deviation results. Finally, Section 5 includes some discussions on the main results of this paper, and the Appendix, i.e. Section 6 contains some technical proofs.

Notations. Throughout this paper, for $x \in \mathbb{R}^+$, $[x]$ denotes the largest integer not greater than x, $f = \mathcal{O}(g)$ means that $|f/g|$ stays bounded, and $f_n = o(1)$ means that $|f_n|$ tends to zero, as $n \to \infty$.

2 Preliminaries

Let us first recall some notations from our recent paper [23]. In particular, we will classify the circuits according to their types. Then, we shall formulate the path expansion formula of the trace precisely.

Let $k \geq 1$ be fixed. For every $0 \leq l \leq [\frac{k}{2}]$, $1 \leq i \leq n-l$, set

$$Q_{l,i}^{\overrightarrow{m}_l, \overrightarrow{n}_l} := \prod_{j=0}^{l-1} (a_{i+j} b_{i+j})^{m_{j+1}} \prod_{j=0}^{l} d_{i+j}^{n_j}, \tag{2.1}$$

where, $\overrightarrow{m}_l = (m_1, m_2, ..., m_l)$ and $\overrightarrow{n}_l = (n_0, n_1, ..., n_l)$. We say that $\overrightarrow{m}_l$ and

$\overrightarrow{n}_l$ are admissible, if $m_j \geq 0$ and $n_h \geq 0$ for every $1 \leq j \leq l$, $0 \leq h \leq l$,

$$2\sum_{j=1}^{l} m_j + \sum_{h=0}^{l} n_h = k, \tag{2.2}$$

and if for some $l > 0$, there exists $1 \leq p \leq l$, such that $m_p = 0$, then $m_j = 0$ and $n_h = 0$ for all $p < j \leq l$, $p \leq h \leq l$. Set

$$\Psi_k := \{(l, \overrightarrow{m}_l, \overrightarrow{n}_l) : 0 \leq l \leq [\frac{k}{2}], \overrightarrow{m}_l \textit{ and } \overrightarrow{n}_l \textit{ are admissible}\}. \tag{2.3}$$

The intuitions of these quantities can be seen as follows. By the tridiagonal structure of (1.1), we note that

$$TrQ_n^k = \sum_{\pi \in \mathscr{C}_n} Q_\pi, \quad Q_\pi := \prod_{j=1}^{k} Q(\pi_j, \pi_{j+1}). \tag{2.4}$$

where $\mathscr{C}_n$ denotes the set of all circuits, i.e.

$$\mathscr{C}_n = \{\pi : (1, \cdots, k) \to (1, \cdots, n) : |\pi_j - \pi_{j+1}| \leq 1, 1 \leq j \leq k, \pi_{k+1} = \pi_1\}.$$

The types of $\mathscr{C}_n$ can be determined by the set Ψ_k and the vertices i, $1 \leq i \leq n$. Indeed, l is the largest length between various vertices in the circuit π, and i is the leftmost vertex in π. The vector $\overrightarrow{m}_l$ determines the circuit $\widehat{\pi}$ consisting of the subindices of off-diagonal entries in Q_π, and the coordinate m_{j+1} is half of the number of edges with the vertices $i + j$ and $i + j + 1$. Similarly, the vector $\overrightarrow{n}_l$ is related to the diagonal entries in Q_π, and the coordinate n_j is the number of loops with the vertex $i + j$. Thus, $Q_{l,i}^{\overrightarrow{m}_l, \overrightarrow{n}_l}$ defined in (2.1) can be viewed as a representative element of the type $(l, \overrightarrow{m}_l, \overrightarrow{n}_l)$ and the leftmost point i.

Let $C_{l,i}^{\overrightarrow{m}_l, \overrightarrow{n}_l}$ denote the number of circuits of the same type $(l, \overrightarrow{m}_l, \overrightarrow{n}_l)$ and the leftmost vertex i. Note that for every $1 \leq i, j \leq n$, $C_{l,i}^{\overrightarrow{m}_l, \overrightarrow{n}_l} = C_{l,j}^{\overrightarrow{m}_l, \overrightarrow{n}_l}$. That is, the number of circuits with the same type $\{l, \overrightarrow{m}_l, \overrightarrow{n}_l\}$, though different leftmost vertices, are also the same. Hence, we can set $C_l^{\overrightarrow{m}_l, \overrightarrow{n}_l} := C_{l,i}^{\overrightarrow{m}_l, \overrightarrow{n}_l}$.

With these notations, the expansion of the trace (2.4) can be reformulated according to the types of the circuits $\mathscr{C}_n$, i.e.

$$TrQ_n^k = \sum_{(l, \overrightarrow{m}_l, \overrightarrow{n}_l) \in \Psi_k} C_l^{\overrightarrow{m}_l, \overrightarrow{n}_l} \sum_{i=1}^{n-l} Q_{l,i}^{\overrightarrow{m}_l, \overrightarrow{n}_l}. \tag{2.5}$$

This formula was recently used in our article [23] for the study of limiting spectral distribution of the random birth-death Q matrix and, as we shall see later, is crucial for the formulations and proofs of the main results in this paper.

We conclude this section by taking TrQ_n^3 and TrQ_n^4 for examples. For TrQ_n^3, in this case $l = 0, 1$, Ψ_k contains the types $(0, (0), (3))$, $(1, (1), (1, 0))$ and $(1, (1), (0, 1))$, and $C_l^{\vec{m}_l, \vec{n}_l}$ is equal to $1, 3, 3$ respectively. Hence, by (2.5) we have

$$TrQ_n^3 = \sum_{i=1}^{n} d_i^3 + 3\sum_{i=1}^{n-1}(a_i b_i)d_i + 3\sum_{i=1}^{n-1}(a_i b_i)d_{i+1}.$$

For TrQ_n^4, $l = 0, 1, 2$, Ψ_k have the types $(0, (0), (4))$, $(1, (1), (2, 0))$, $(1, (1), (1, 1))$, $(1, (1), (0, 2))$, $(1, (2), (0, 0))$ and $(1, (1, 1), (0, 0, 0))$, and the corresponding $C_l^{\vec{m}_l, \vec{n}_l}$ are $1, 4, 4, 4, 2$ and 4 respectively. Thus, it follows that

$$TrQ_n^4 = \sum_{i=1}^{n} d_i^4 + \sum_{i=1}^{n-1} a_i b_i(4d_i^2 + 4d_i d_{i+1} + 4d_{i+1}^2 + 2a_i b_i) + 4\sum_{i=1}^{n-2} a_i b_i a_{i+1} b_{i+1}.$$

3 Gaussian fluctuations

In this section, the main results Theorem 3.2 and 3.3 are formulated in Subsection 3.1, and then they are proved in Subsection 3.2 and 3.3 respectively. Moreover, several applications are also given in Subsection 3.4.

3.1 Main results

Let m be a fixed nonnegative integer. The random variables $\{X_i\}$ are said to be m-dependent, if for any positive integers i and j with $j - i > m$, X_j is independent of the σ-field generated by $\{X_h, 1 \leq h \leq i\}$. In particular, 0-dependence is equivalent to independence.

Motivated by the physical models and random matrix models mentioned in Section 1, we introduce the assumptions $(H.1)$ and $(H.2)$ below.

(H.1) In the symmetric case (i.e. $a_i = b_i$ for all $i \geq 1$), the off-diagonal entries $\{a_i\}$ are independent, and the diagonal entries $\{d_i\}$ satisfy that $d_i = f(a_{i-1}, a_i)$ with f a continuous function on $\mathbb{R}^2$ or $\{a_i, d_i\}$ are all independent.

In the non-symmetric case, the random vectors $\{(a_{i-1}, d_i, b_i)\}$ are independent.

As regards the asymptotic behavior of the entries in (1.1), we assume that

(H.2) Let α and ε be two nonnegative constants, $0 \le \varepsilon \le \alpha$. There exist constants a, d, b and random variables η_i, ζ_i, ξ_i, η, ζ, ξ, such that

$$i^{-\alpha}(a_{i-1}, d_i, b_i) = (a, d, b) + i^{-\varepsilon}(\eta_{i-1}, \zeta_i, \xi_i), \tag{3.1}$$

where η_{i-1}, ζ_i and ξ_i satisfy

$$(\eta_{i-1}, \zeta_i, \xi_i) \xrightarrow{d} (\eta, \zeta, \xi), \tag{3.2}$$

" $\xrightarrow{d}$ " means convergence in distribution, and for each $k \ge 1$,

$$\sup_{i \ge 1} \mathbb{E}(\eta_i^{2k} + \zeta_i^{2k} + \xi_i^{2k}) < \infty. \tag{3.3}$$

Remark 3.1 (i). *By Hölder's inequality,* (3.3) *implies that all moments of* η_n, ζ_n, ξ_n *are finite.*

(ii). *Assumption* $(H.1)$ *allows to treat the tridiagonal random matrix* (1.1) *with non-independent entries. The conditions on* d_i *in the symmetric case is mainly motivated by the random birth-death Markov kernel and the random birth-death* Q *matrix, where* $f(a_{i-1}, a_i)$ *is of the form* $1 - (a_{i-1} + a_i)$ *and* $-(a_{i-1} + a_i)$ *respectively.*

(iii). *In the case that the entries of* (1.1) *are independent and identically distributed (i.i.d.) and have all moments finite, the assumptions* $(H.1)$ *and* $(H.2)$ *are obviously verified.*

(iv). *In Assumption* $(H.2)$ *with* $\alpha > 0$, *when* $0 < \varepsilon < \alpha$ *(resp.* $\varepsilon = 0$*), the entries convergence in distribution to degenerate (resp. non-degenerate) random variables. The degenerate case actually includes the* β*-Hermite ensemble. See Subsection 3.4 for more details.*

The main results in this section are formulated below. Taking into account the complicated formulations of covariances in the case $\varepsilon > 0$, we shall consider the case $\varepsilon = 0$ and $0 < \varepsilon \le \alpha$ in Theorem 3.2 and 3.3 respectively, to make the structure of the variance and covariances clean.

Theorem 3.2 *Assume (H.1) and (H.2) with $\varepsilon = 0$.*

(i). For each $k \geq 1$, let $m_k = [\frac{k}{2}]$ (resp. $[\frac{k}{2}] + 1$) in the non-symmetric (resp. symmetric) case. Set $Tr\widetilde{Q_n^k} := TrQ_n^k - \mathbb{E}TrQ_n^k$. Then,

$$n^{-(\alpha k+\frac{1}{2})}Tr\widetilde{Q_n^k} \xrightarrow{d} N(0, D_k). \tag{3.4}$$

Here

$$D_k = \frac{1}{2\alpha k+1}\left[Var(Z_{k,1}) + 2\sum_{j=1}^{m_k} Cov(Z_{k,1}, Z_{k,1+j})\right], \tag{3.5}$$

and for each $1 \leq i \leq m_k + 1$,

$$Z_{k,i} = \sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k} C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l} \prod_{j=0}^{l-1}(\widetilde{\eta}_{i+j}\widetilde{\xi}_{i+j})^{m_{j+1}} \prod_{j=0}^{l}(\widetilde{\zeta}_{i+j})^{n_j}, \tag{3.6}$$

where $\widetilde{\eta}_i, \widetilde{\zeta}_i, \widetilde{\xi}_i$ have the same distributions as those of $a + \eta, d + \zeta$ and $b + \xi$ respectively, and they satisfy Assumption (H.1) with all a_i, d_i, b_i replaced by $\widetilde{\eta}_i, \widetilde{\zeta}_i$ and $\widetilde{\xi}_i$ respectively.

(ii) Given $k_1, \cdots, k_r \geq 1, r \geq 1$, set $m_{ij} = \max\{m_{k_i}, m_{k_j}\}$. We have

$$n^{-\frac{1}{2}}\ (n^{-\alpha k_1}Tr\widetilde{Q_n^{k_1}}, \cdots, n^{-\alpha k_r}Tr\widetilde{Q_n^{k_r}}) \xrightarrow{d} \Phi_\Lambda, \tag{3.7}$$

where Φ_Λ is a r-dimensional normal distribution with mean zero and covariance matrix Λ, for $1 \leq i, j \leq r$,

$$\begin{aligned}\Lambda(i,j) =& \frac{1}{\alpha(k_i+k_j)+1}\Bigg[Cov(Z_{k_i,1}, Z_{k_j,1}) \\ &+ \sum_{h=1}^{m_{ij}}\big(Cov(Z_{k_i,1}, Z_{k_j,1+h}) + Cov(Z_{k_i,1+h}, Z_{k_j,1})\big)\Bigg],\end{aligned} \tag{3.8}$$

where $Z_{k_i,h}$ and $Z_{k_j,h}$, $1 \leq h \leq m_{ij} + 1$, are defined as in (3.6).

Theorem 3.3 *Assume (H.1) and (H.2) with $0 < \varepsilon \leq \alpha$. Let $Tr\widetilde{Q_n^k}, m_k$ and m_{ij} be as in Theorem 3.2.*

(i). For each $k \geq 1$, $0 \leq j \leq m_k$, there exists the limit

$$\sigma_k(1+j, 1) := \lim_{n\to\infty} n^{-2(\alpha k-\varepsilon)}Cov(X_{k,n}, X_{k,n-j}), \tag{3.9}$$

and we have

$$n^{-(\alpha k+\frac{1}{2}-\varepsilon)}Tr\widetilde{Q_n^k} \xrightarrow{d} N(0, D_k),$$

where

$$D_k = \frac{1}{2\alpha k+1-2\varepsilon}(\sigma_k(1,1)+2\sum_{j=1}^{m_k}\sigma_k(1+j,1)).$$

(ii). Given $k_1, \cdots, k_r \geq 1$, $r \geq 1$. For every $1 \leq i, j \leq r$ and $0 \leq h \leq m_{ij}$, there exists the limit

$$\sigma_{k_i,k_j}(1+h,1) := \lim_{n\to\infty} n^{-(\alpha(k_i+k_j)-2\varepsilon)}Cov(X_{k_i,n}, X_{k_j,n-h}), \tag{3.10}$$

and we have

$$n^{-\frac{1}{2}+\varepsilon}(n^{-\alpha k_1}Tr\widetilde{Q_n^{k_1}}, \cdots, n^{-\alpha k_r}Tr\widetilde{Q_n^{k_r}}) \xrightarrow{d} \Phi_\Lambda,$$

where Φ_Λ is the r-dimensional normal distribution as in Theorem 3.2, but with the covariances $\Lambda(i,j)$ defined by

$$\begin{aligned}\Lambda(i,j) =& \frac{1}{\alpha(k_i+k_j)+1-2\varepsilon}\Big[\sigma_{k_i,k_j}(1,1) \\ &+ \sum_{h=1}^{m_{ij}}\big(\sigma_{k_i,k_j}(1,1+h)+\sigma_{k_i,k_j}(1+h,1)\big)\Big],\end{aligned} \tag{3.11}$$

where $\sigma_{k_i,k_j}(1,1+h) = \sigma_{k_j,k_i}(1+h,1)$.

Remark 3.4 *Actually, in the case $0 < \varepsilon \leq \alpha$, $\sigma_k(1+j,1)$ and $\sigma_{k_i,k_j}(1+h,1)$ can be calculated explicitly in terms of the covariances of η_n, ξ_n and ζ_n. Since the formulations are complicated, we omit them in the statement of Theorem 3.3. Concrete calculations are shown in Corollary 3.6 below for the symmetric tridiagonal random matrix motivated by [29] and the β-Hermite ensemble.*

3.2 Proof of Theorem 3.2

The key observation for the proof is that, the path expansion formula of the trace (2.5) indicates that the fluctuation of TrQ_n^k is approximately the same as that of

$$\sum_{(l,\vec{m}_l,\vec{n}_l)\in\Psi_k} C_l^{\vec{m}_l,\vec{n}_l} \sum_{i=1}^n Q_{l,i}^{\vec{m}_l,\vec{n}_l} = \sum_{i=1}^n X_{k,i},$$

where

$$\begin{aligned}X_{k,i} &:= \sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k} C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l} Q_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l} \\ &= \sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k} C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l} \prod_{j=0}^{l-1}(a_{i+j}b_{i+j})^{m_{j+1}} \prod_{j=0}^{l} d_{i+j}^{n_j}. \end{aligned} \tag{3.12}$$

$\{X_{k,i}\}_{i\geq 1}$ is m_k-dependent with m_k defined as in Theorem 3.2, due to Assumption $(H.1)$ and the finite width of band in the tridiagonal random matrix (1.1). With this new point of view, we can expect the Gaussian fluctuation of TrQ_n^k, inspired by the fact that m-dependent random variables are approximated normally distributed in the stationary case (see e.g. [24, 9, 19]). On the technical level, in order to deal with the non-stationary case $\alpha > 0$ in Assumption $(H.2)$, we will employ the standard blocking arguments to separate the sum $\sum_{i=1}^n X_{k,i}$ into independent blocks and small ones.

Proof of Theorem 3.2. (i). We first note that

$$\mathbb{E}\Big|n^{-(\alpha k+\frac{1}{2})}(Tr\widetilde{Q_n^k} - \sum_{i=1}^n \widetilde{X}_{k,i})\Big|^2 \to 0, \tag{3.13}$$

where $\widetilde{X}_{k,i} = X_{k,i} - \mathbb{E}X_{k,i}$. (See the Appendix for the proof.)

Thus, we only need to consider the fluctuation of $\sum_{i=1}^n \widetilde{X}_{k,i}$, which is actually the sum of m_k-dependent random variables.

Let $\kappa < \frac{1}{4}$, $n' = [n^\kappa]$, $p = [\frac{n}{n'}]$, $r = n - pn'$. Set $\widetilde{Y}_{n,i} = n^{-\alpha k}\widetilde{X}_{k,i}$. For $1 \leq i \leq p$, let $\widetilde{U}_{n,i} := \sum\limits_{j=(i-1)n'+1}^{in'-m_k} \widetilde{Y}_{n,j}$. For $1 \leq i \leq p-1$, set $\widetilde{Z}_{n,i} = \sum\limits_{j=in'-m_k+1}^{in'} \widetilde{Y}_{n,j}$, $\widetilde{Z}_{n,p} = \sum\limits_{j=pn'-m_k+1}^{n} \widetilde{Y}_{n,j}$ and $\widetilde{T}_n = \sum\limits_{i=1}^{p} \widetilde{Z}_{n,i}$. Note that, $\{\widetilde{U}_{n,i}\}_{1\leq i\leq p}$ are independent, and so are $\{\widetilde{Z}_{n,i}\}_{1\leq i\leq p}$, due to the m_k-dependence of $\{X_{k,i}\}$. Moreover,

$$n^{-(\alpha k+\frac{1}{2})}\sum_{i=1}^n \widetilde{X}_{k,i} = n^{-\frac{1}{2}}\sum_{i=1}^n \widetilde{Y}_{n,i} = n^{-\frac{1}{2}}\sum_{i=1}^p \widetilde{U}_{n,i} + n^{-\frac{1}{2}}\widetilde{T}_n.$$

Let us show that, as $n \to \infty$,

$$n^{-\frac{1}{2}}\widetilde{T}_n \xrightarrow{d} 0. \tag{3.14}$$

Indeed, by the independence of $\{\widetilde{Z}_{n,i}\}$,

$$\mathbb{E}(n^{-\frac{1}{2}}\widetilde{T}_n)^2 = \frac{1}{n}\mathbb{E}(\sum_{i=1}^{p}\widetilde{Z}_{n,i})^2 = \frac{1}{n}\sum_{i=1}^{p}\mathbb{E}\widetilde{Z}_{n,i}^2.$$

Then, it follows from the boundedness of moments in (3.3) that

$$\mathbb{E}(n^{-\frac{1}{2}}\widetilde{T}_n)^2 = \mathcal{O}(\frac{pm_k^2}{n}) + \mathcal{O}(\frac{(n-pn'+m_k)^2}{n}) \to 0, \tag{3.15}$$

which yields (3.14), as claimed.

In view of (3.14) and $pn'/n \to 1$, the proof of Theorem 3.2 now reduces to proving that

$$\frac{1}{\sqrt{pn'}}\sum_{i=1}^{p}\widetilde{U}_{n,i} \xrightarrow{d} N(0, D_k). \tag{3.16}$$

For this purpose, let us first treat the limit behavior of the variance. Using the notations as in [24], we set

$$A_{n,i} := \mathbb{E}\widetilde{Y}_{n,i+m_k}^2 + 2\sum_{j=1}^{m_k}\mathbb{E}\widetilde{Y}_{n,i+m_k}\widetilde{Y}_{n,i+m_k-j}. \tag{3.17}$$

By the independence of $\{\widetilde{U}_{n,i}\}_{1\le i\le p}$ and the m_k-dependence of $\{\widetilde{Y}_{n,i}\}_{1\le i\le n}$, straightforward computations show that

$$\begin{aligned}
&\mathbb{E}(\frac{1}{\sqrt{pn'}}\sum_{i=1}^{p}\widetilde{U}_{n,i})^2 = \frac{1}{pn'}\sum_{i=1}^{p}\mathbb{E}(\widetilde{U}_{n,i})^2 \\
=&\frac{1}{pn'}\sum_{i=1}^{p}\left[\mathbb{E}\left(\sum_{h=1}^{m_k}\widetilde{Y}_{n,(i-1)n'+h}\right)^2 + \sum_{h=1}^{n'-2m_k}A_{n,(i-1)n'+h}\right] \\
=&\mathcal{O}(\frac{m_k^2}{n'}) + \frac{1}{pn'}\sum_{i=1}^{p}\sum_{h=1}^{n'-2m_k}A_{n,(i-1)n'+h},
\end{aligned} \tag{3.18}$$

where the last step is due to (3.3).

Set $n_{i,h} := (i-1)n' + h + m_k$. In order to take the limit $p, n' \to \infty$, we note from (3.17) and the definition of $\widetilde{Y}_{n,i}$ that

$$
\begin{aligned}
&\frac{1}{pn'}\sum_{i=1}^{p}\sum_{h=1}^{n'-2m_k} A_{n,(i-1)n'+h} \\
=&\frac{1}{pn'}\sum_{i=1}^{p}\sum_{h=1}^{n'-2m_k}\left[\mathbb{E}\widetilde{Y}^2_{n,n_{i,h}} + 2\sum_{j=1}^{m_k}\mathbb{E}\widetilde{Y}_{n,n_{i,h}}\widetilde{Y}_{n,n_{i,h}-j}\right] \\
=&\frac{1}{pn'}\sum_{i=1}^{p}\sum_{h=1}^{n'-2m_k}(\frac{n_{i,h}}{n})^{2\alpha k} Var(\frac{X_{k,n_{i,h}}}{n_{i,h}^{\alpha k}}) \\
&+\frac{2}{pn'}\sum_{j=1}^{m_k}\sum_{i=1}^{p}\sum_{h=1}^{n'-2m_k}(\frac{n_{i,h}}{n})^{\alpha k}(\frac{n_{i,h}-j}{n})^{\alpha k} Cov(\frac{X_{k,n_{i,h}}}{n_{i,h}^{\alpha k}}, \frac{X_{k,n_{i,h}-j}}{(n_{i,h}-j)^{\alpha k}}). \quad (3.19)
\end{aligned}
$$

It is not difficult to see that, for every $0 \le j \le m_k$, as $p, n' \to \infty$,

$$\frac{1}{pn'}\sum_{i=1}^{p}\sum_{h=1}^{n'-2m_k}(\frac{n_{i,h}}{n})^{\alpha k}(\frac{n_{i,h}-j}{n})^{\alpha k} \to \frac{1}{2\alpha k+1}. \quad (3.20)$$

Moreover, we have that as $n(i,h) \to \infty$,

$$Cov(\frac{X_{k,n_{i,h}}}{n_{i,h}^{\alpha k}}, \frac{X_{k,n_{i,h}-j}}{(n_{i,h}-j)^{\alpha k}}) \to Cov(Z_{k,1}, Z_{k,1+j}), \quad (3.21)$$

where $Z_{k,i}$, $1 \le i \le m_k + 1$, are defined as in (3.6). (See the Appendix for the proof.)

Thus, plugging (3.20) and (3.21) into (3.19), we obtain that as $p, n' \to \infty$,

$$\frac{1}{pn'}\sum_{i=1}^{p}\sum_{h=1}^{n'-2m_k} A_{n,(i-1)n'+h} \to D_k \quad (3.22)$$

with D_k defined as in (3.5), which together with (3.18) yields that

$$\mathbb{E}(\frac{1}{\sqrt{pn'}}\sum_{i=1}^{p}\widetilde{U}_{n,i})^2 \to D_k. \quad (3.23)$$

As regards the third moments of $\widetilde{U}_{n,i}/\sqrt{n'}$ in (3.16), since $\mathbb{E}|\widetilde{Y}_{n,j}|^3 = \mathcal{O}(1)$ and $n' = o(p^{\frac{1}{3}})$, we have for $1 \le i \le p$,

$$\begin{aligned}
\mathbb{E}\left|\frac{1}{\sqrt{n'}}\widetilde{U}_{n,i}\right|^3 =&(n')^{-\frac{3}{2}}\mathbb{E}\left|\sum_{j=(i-1)n'+1}^{in'-m_k}\widetilde{Y}_{n,j}\right|^3 \\
=&\mathcal{O}((n')^{-\frac{3}{2}}(n'-m_k)^3) \\
=&\mathcal{O}((n')^{\frac{3}{2}}) = o(p^{\frac{1}{2}}).
\end{aligned} \tag{3.24}$$

Consequently, combining (3.23) and (3.24), and applying Lyapunov's central limit theorem (see [24, Appendix], see also [9, Theorem 7.1.2]), we obtain (3.16) and complete the proof of the first part (i).

(ii). In order to prove (3.7), we just need to show that for all $z_1, \cdots, z_r \in \mathbb{R}$, the linear combination $n^{-1/2}\sum_{j=1}^r z_j n^{-\alpha k_j} \widetilde{TrQ_n^{k_j}}$ is normally distributed in the limit.

To this end, as in the proof of (3.13), the limit behavior is the same as that of

$$n^{-\frac{1}{2}}\sum_{j=1}^{r} z_j n^{-\alpha k_j}\sum_{i=1}^{n} X_{k_j,i} = n^{-\frac{1}{2}}\sum_{i=1}^{n}\sum_{j=1}^{r} z_j n^{-\alpha k_j} X_{k_j,i}, \tag{3.25}$$

where $X_{k_j,i}$ is defined as in (3.12) with k_j replacing k.

Since $\{X_{k_j,i}\}_{i\ge1}$ is m_{k_j}-dependent for each $1 \le j \le r$, $\{\sum_{j=1}^r z_j n^{-\alpha k_j} X_{k_j,i}\}_{i\ge1}$ is M-dependent with $M = \max_{1\le j\le r} m_{k_j}$. Thus, as in the preceding proof of the first assertion (i), we deduce that $n^{-1/2}\sum_{i=1}^n\sum_{j=1}^r z_j n^{-\alpha k_j} X_{k_j,i}$ is normally distributed in the limit, which consequently implies the Gaussian fluctuation of the random vectors $n^{-\frac{1}{2}}(n^{-\alpha k_1}\widetilde{TrQ_n^{k_1}}, \cdots, n^{-\alpha k_r}\widetilde{TrQ_n^{k_r}})$, $n \ge 1$.

It remains to compute the covariances $\Lambda(i,j)$. Let $m_{ij} = \max\{m_{k_i}, m_{k_j}\}$. Then, $\{X_{k_i,q}, X_{k_j,q}\}_{q\ge1}$ is m_{ij}-dependent. Set $\widetilde{Y}_{k_i,n,q} := n^{-\alpha k_i}\widetilde{X}_{k_i,q}$, and

$$\begin{aligned}
B_{n,q} :=&\mathbb{E}\widetilde{Y}_{k_i,n,q+m_{ij}}\widetilde{Y}_{k_j,n,q+m_{ij}} \\
&+\sum_{h=1}^{m_{ij}}(\mathbb{E}\widetilde{Y}_{k_i,n,q+m_{ij}-h}\widetilde{Y}_{k_j,n,q+m_{ij}} + \mathbb{E}\widetilde{Y}_{k_i,n,q+m_{ij}}\widetilde{Y}_{k_j,n,q+m_{ij}-h}).
\end{aligned}$$

Straightforward computations show that

$$
\begin{aligned}
&n^{-(\alpha(k_i+k_j)+1)}Cov(\sum_{q=1}^{n} X_{k_i,q}, \sum_{q=1}^{n} X_{k_j,q}) \\
=&n^{-1}\mathbb{E}(\sum_{q=1}^{m_{ij}} \widetilde{Y}_{k_i,n,q})(\sum_{q=1}^{m_{ij}} \widetilde{Y}_{k_j,n,q}) + n^{-1}\sum_{q=1}^{n-m_{ij}} B_{n,q} \\
=&\mathcal{O}(n^{-1}) + n^{-1}\sum_{q=1}^{n-m_{ij}} \mathbb{E}\widetilde{Y}_{k_i,n,q+m_{ij}}\widetilde{Y}_{k_j,n,q+m_{ij}} \\
&+ n^{-1}\sum_{h=1}^{m_{ij}}\sum_{q=1}^{n-m_{ij}} (\mathbb{E}\widetilde{Y}_{k_i,n,q+m_{ij}-h}\widetilde{Y}_{k_j,n,q+m_{ij}} + \mathbb{E}\widetilde{Y}_{k_i,n,q+m_{ij}}\widetilde{Y}_{k_j,n,q+m_{ij}-h}).
\end{aligned}
\tag{3.26}
$$

As in the proof of (3.22), we deduce that for each $0 \leq h \leq m_{ij}$, as $n \to \infty$,

$$
\frac{1}{n}\sum_{q=1}^{n-m_{ij}} (\frac{q+m_{ij}-h}{n})^{\alpha k_i}(\frac{q+m_{ij}}{n})^{\alpha k_j} \to \frac{1}{\alpha(k_i+k_j)+1},
$$

and as $q \to \infty$,

$$
Cov(\frac{X_{k_i,q+m_{ij}-h}}{(q+m_{ij}-h)^{\alpha k_i}}, \frac{X_{k_j,q+m_{ij}}}{(q+m_{ij})^{\alpha k_j}}) \to Cov(Z_{k_i,1}, Z_{k_j,1+h}).
$$

Hence,

$$
n^{-1}\sum_{q=1}^{n-m_{ij}} \mathbb{E}(\widetilde{Y}_{k_i,n,q+m_{ij}-h}\widetilde{Y}_{k_j,n,q+m_{ij}}) \to \frac{1}{\alpha(k_i+k_j)+1}Cov(Z_{k_i,1}, Z_{k_j,1+h}).
\tag{3.27}
$$

Similarly,

$$
n^{-1}\sum_{q=1}^{n-m_{ij}} \mathbb{E}(\widetilde{Y}_{k_i,n,q+m_{ij}}, \widetilde{Y}_{k_j,n,q+m_{ij}-h}) \to \frac{1}{\alpha(k_i+k_j)+1}Cov(Z_{k_i,1+h}, Z_{k_j,1}).
\tag{3.28}
$$

Therefore, it follows from (3.26)-(3.28) that, as $n \to \infty$,

$$
n^{-(\alpha(k_i+k_j)+1)}Cov(\sum_{q=1}^{n} X_{k_i,q}, \sum_{q=1}^{n} X_{k_j,q}) \to \Lambda(i,j).
$$

where $\Lambda(i,j)$ is defined as in (3.8), thereby completing the proof. $\square$

3.3 Proof of Theorem 3.3

Proof of Theorem 3.3. The proof is similar to that of Theorem 3.2, we only need to prove that, with the additional scaling $n^{2\varepsilon}$, the limits (3.9) and (3.10) exist. Below, we show the existence of the limit (3.9), the proof of (3.10) follows analogously.

Taking into account the definition of $X_{k,i}$ in (3.12), we need to show the asymptotic estimate of $\prod_{j=0}^{l-1}(a_{n+j}b_{n+j})^{m_{j+1}}\prod_{j=0}^{l} d_{n+j}^{n_j}$.

Indeed, by Assumption $(H.2)$, for each $0 \le j \le l-1$,

$$\frac{a_{n+j}b_{n+j}}{n^{2\alpha}} = ab + b\frac{\eta_{n+j}}{n^{\varepsilon}} + a\frac{\xi_{n+j}}{n^{\varepsilon}} + \mathcal{O}(\frac{1}{n^{2\varepsilon}}).$$

Here, with a slight abuse of notation, $\mathcal{O}(n^{-2\varepsilon})$ stands for the term of order $n^{-2\varepsilon}$ after taking the expectation. Hence, for $m_{j+1} \ge 1$,

$$\begin{aligned}&(\frac{a_{n+j}b_{n+j}}{n^{2\alpha}})^{m_{j+1}}\\ =&(ab)^{m_{j+1}} + (ab)^{m_{j+1}-1}m_{j+1}\left(b\frac{\eta_{n+j}}{n^{\varepsilon}} + a\frac{\xi_{n+j}}{n^{\varepsilon}}\right) + \mathcal{O}(\frac{1}{n^{2\varepsilon}}).\end{aligned}$$

Hence, setting $|\overrightarrow{m}_l| = \sum_{j=1}^{l} m_j$, we derive that for $|\overrightarrow{m}_l| \ge 1$,

$$\begin{aligned}&\prod_{j=0}^{l-1}(\frac{a_{n+j}b_{n+j}}{n^{2\alpha}})^{m_{j+1}}\\ =&(ab)^{|\overrightarrow{m}_l|-1}[ab + a\sum_{j=1}^{l} m_j\frac{\xi_{n+j-1}}{n^{\varepsilon}} + b\sum_{j=1}^{l} m_j\frac{\eta_{n+j-1}}{n^{\varepsilon}}] + \mathcal{O}(\frac{1}{n^{2\varepsilon}}).\end{aligned} \tag{3.29}$$

Similarly, setting $|\overrightarrow{n}_l| = \sum_{j=0}^{l} n_j$, we have that for $|\overrightarrow{n}_l| \ge 1$,

$$\begin{aligned}\prod_{j=0}^{l}(\frac{d_{n+j}}{n^{\alpha}})^{n_j} =&\prod_{j=0}^{l}\left(d^{n_j} + d^{n_j-1}n_j\frac{\zeta_{n+j}}{n^{\varepsilon}}\right) + \mathcal{O}(\frac{1}{n^{2\varepsilon}})\\ =&d^{|\overrightarrow{n}_l|-1}[d + \sum_{j=0}^{l} n_j\frac{\zeta_{n+j}}{n^{\varepsilon}}] + \mathcal{O}(\frac{1}{n^{2\varepsilon}}).\end{aligned} \tag{3.30}$$

Thus, (3.29) and (3.30) yield that for $|\overrightarrow{m}_l| \geq 1$, $|\overrightarrow{n}_l| \geq 1$,

$$\begin{aligned}
&\prod_{j=0}^{l-1}(\frac{a_{n+j}b_{n+j}}{n^{2\alpha}})^{m_{j+1}} \prod_{j=0}^{l}(\frac{d_{n+j}}{n^{\alpha}})^{n_j} \\
=&(ab)^{|\overrightarrow{m}_l|-1}d^{|\overrightarrow{n}_l|-1}\Bigg[abd + \sum_{j=1}^{l}\Bigg(adm_j\xi_{n+j-1} + bdm_j\eta_{n+j-1} \\
&+ abn_j\zeta_{n+j}\Bigg)n^{-\varepsilon} + abn_0\zeta_n n^{-\varepsilon}\Bigg] + \mathcal{O}(n^{-2\varepsilon}).
\end{aligned} \tag{3.31}$$

Moreover, it is easy to see that, for $|\overrightarrow{m}_l| = 0$,

$$(\frac{d_n}{n^{\alpha}})^k = d^{k-1}(d + k\frac{\zeta_n}{n^{\varepsilon}}) + \mathcal{O}(\frac{1}{n^{2\varepsilon}}), \tag{3.32}$$

and for $|\overrightarrow{n}_l| = 0$,

$$\begin{aligned}
&\prod_{j=0}^{l-1}(\frac{a_{n+j}b_{n+j}}{n^{2\alpha}})^{m_{j+1}} \\
=&(ab)^{\frac{k}{2}-1}[ab + a\sum_{j=1}^{l}m_j\frac{\xi_{n+j-1}}{n^{\varepsilon}} + b\sum_{j=1}^{l}m_j\frac{\eta_{n+j-1}}{n^{\varepsilon}}] + \mathcal{O}(\frac{1}{n^{2\varepsilon}}).
\end{aligned} \tag{3.33}$$

Plugging (3.12), (3.31)-(3.33) into the right hand side of (3.9), we consequently deduce that the limit (3.9) exists, which actually can be computed explicitly in terms of the covariances of η_n, ξ_n and ζ_n. For simplicity, we omit it here. The proof is complete. $\square$

3.4 Applications

Let us start with the independent identically distributed (i.i.d.) case, which is a direct consequence of Theorem 3.2 .

Corollary 3.5 *(The case $\alpha = 0$.) In the symmetric case, assume that $\{a_i\}$ is a sequence of i.i.d. random variables, $\{d_i\}$ satisfy that $d_i = f(a_{i-1}, a_i)$ with f a continuous function, or $\{d_i\}$ are i.i.d random variables and independent of $\{a_i\}$. In the non-symmetric case, assume that $\{(a_{i-1}, d_i, b_i)\}$ is a sequence of i.i.d random vectors. Suppose also that in both cases the entries have all*

moments finite. Then, the assertions in Theorem 3.2 hold. In particular, the variance D_k is given by

$$D_k = Var(X_{k,2}) + 2\sum_{j=1}^{m_k} Cov(X_{k,2}, X_{k,2+j}), \tag{3.34}$$

with m_k defined as in Theorem 3.2 and $X_{k,i}$ as in $(3.12)^3$, $2 \le i \le m_k + 2$, *and the covariances $\Lambda(i,j)$ are given by*

$$\Lambda(i,j) = Cov(X_{k_i,2}, X_{k_j,2}) + \sum_{h=1}^{m_{ij}} (Cov(X_{k_i,2}, X_{k_j,2+h}) + Cov(X_{k_i,2+h}, X_{k_j,2})),$$

where $m_{ij} = \max\{m_{k_i}, m_{k_j}\}$.

Corollary 3.5 have several applications to the physical models and random matrix models mentioned in Section 1.

Anderson model ([28, 7]). When $a_i = b_i = -1$, and $\{d_i\}$ are i.i.d random variables with all moments finite, this tridiagonal random matrix (1.1) is referred to the Anderson model in the physical literature, restricted to the bound domain $\Lambda_n = \{1, \cdots, n\}$. Write $Q_n = Q_{n,0} + V_n$, where $V_n = diag(d_1, \cdots, d_n)$. $Q_{n,0}$ represents the finite difference operator, a discrete analogue of the one dimensional Schrödinger operator, and V_n represents the operator of multiplication by the random field d_i, $1 \le i \le n$. For this model, the conditions in Corollary 3.5 are verified. Hence, its traces of powers TrQ_n^k, $k \ge 1$, are normally distributed in the limit.

Hatano-Nelson model ([21]). When $a_i/b_i > 0$ and $\{(a_{i-1}, d_i, b_i)\}$ is a sequence of i.i.d. random vectors with all moments finite, this tridiagonal random matrix is motivated by the non-Hermitian quantum mechanics of Hatano and Nelson (see [21] and references therein). In this case, it follows from Corollary 3.5 that the traces are approximated normally distributed.

Random birth-death Markov kernel ([5, 8]). This tridiagonal random matrix arises in the random walks with random environment in chain graphs. Given the chain graph $G = (V, E)$, $V = \{1, \cdots, n\}$, $E = \{(i,j), |i-j| \le 1\}$.

[3] The subindices of $X_{k,i}$ here starts with 2, since, e.g. in the case $d_i = f(a_{i-1}, a_i)$, $i \ge 1$, the distribution of $d_1 (= f(0, a_1))$ is different from that of $d_i = f(a_{i-1}, a_i)$, $i \ge 2$.

Consider the random birth-death Markov kernel Q_n, with the state space V and the entries satisfying $a_i, b_i \in (0,1]$, $d_i \in [0,1]$, and $a_{i-1} + d_i + b_i = 1$. One concrete example is the random conductance model, in which case $b_1 = a_{n-1} = 1$, $b_i = 1 - a_{i-1} = U_{i,i+1}/(U_{i,i+1} + U_{i,i-1})$, and $\{U_{i,i+1}\}$ are i.i.d. positive random variables. Another example is that $b_1 = a_{n-1} = 1$, $b_i = 1 - a_{i-1} = V_i$, and $\{V_i\}$ are i.i.d. random variables on $[0,1]$. We refer to [5] for more detail discussions and for the study of the limiting spectral distribution of such model in the ergodic environment. Here, in the i.i.d. environment, i.e. $\{(a_{i-1}, d_i, b_i)\}$ is a sequence of i.i.d. random vectors with all moments finite, we deduce from Corollary 3.5 that the traces are asymptotically normally distributed.

Random birth-death Q matrix ([23]). This model is motivated from the infinitesimal generator of continuous-time Markov process in the chain graph. In this case, $a_i, b_i > 0$ and $d_i = -(a_{i-1} + b_i)$. In [23] we proved the existence of the limiting spectral distribution in the strictly stationary ergodic case (see [23, Theorem 1]). Here, in the case that $\{a_i\}$ and $\{b_i\}$ are two sequences of i.i.d random variables, $\{a_i\}$ is independent of $\{b_i\}$ (or $a_i = b_i$, $i \geq 1$), and they have all moments finite, Corollary 3.5 implies the Gaussian fluctuations of the traces.

Now, we formulate and prove Corollary 3.6 below for the case $\alpha > 0$, which is mainly motivated by the symmetric tridiagonal random matrix studied in [29] and includes, in particular, the well-known β-Hermite ensemble.

For simplicity, we will regard the $l+1$-dimensional vector $\overrightarrow{n}_l$ in Ψ_k in (2.3) as the vector in the larger space $\mathbb{R}^{[\frac{k}{2}]+1}$, by just adding zeros to the remaining $[\frac{k}{2}] - l$ coordinates, and for each $0 \leq q \leq [\frac{k}{2}]$, $\overrightarrow{e}_q$ denotes the q-th normal basis in $\mathbb{R}^{[\frac{k}{2}]+1}$, namely, the q-th coordinate of $\overrightarrow{e}_q$ is 1 and the others are zeros.

Corollary 3.6 *(The case $\alpha > 0$.) Consider the symmetric tridiagonal random matrix* (1.1)*, i.e. $a_i = b_i$, $i \geq 1$. Assume that in* $(H.1)$ *all a_i and d_i are independent. Assume also* $(H.2)$ *with $0 < \varepsilon \leq \alpha$ and $\sup_{n\geq 1} \mathbb{E}|d_n|^k < \infty$ for any $k \geq 1$. Then, the assertions in Theorem 3.3 hold. In particular, the*

covariances $\Lambda(i,j)$ *are given by*

$$\Lambda(i,j)=\begin{cases} \frac{a^{k_i+k_j-2}Var(\eta)}{\alpha(k_i+k_j)+1-2\varepsilon}k_ik_j\binom{k_i}{k_i/2}\binom{k_j}{k_j/2}, & \text{if } k_i,k_j \text{ even;}\\ \frac{a^{k_i+k_j-2}Var(\zeta)}{\alpha(k_i+k_j)+1-2\alpha}k_ik_j\binom{k_i-1}{(k_i-1)/2}\binom{k_j-1}{(k_j-1)/2}, & \text{if } \varepsilon=\alpha,\ k_i,k_j \text{ odd ;}\\ 0, & \text{otherwise.}\end{cases} \tag{3.35}$$

Remark 3.7 *This result coincides with Theorem* 3 *in [29],*[4] *under the conditions considered here. However, the proof presented below is quite different from that in [29], it is analytic and mainly based on Theorem 3.3.*

Proof. First, by the condition on d_n, we note that in Assumption $(H.2)$ $d=0$, $\zeta_n=n^{-(\alpha-\varepsilon)}d_n$, and the limit $\zeta=0$ when $\varepsilon<\alpha$. Below we shall discuss the case k is even or odd respectively.

When k is even, (2.2) indicates that $|\overrightarrow{n}_l|:=\sum_{h=0}^{l}n_h$ is also even. Since $d=0$, from (3.31)-(3.33) we see that the main contribution comes from the case $|\overrightarrow{n}_l|=0$, i.e. there is no loop in the circuit. Hence, it follows from (3.12) and (3.33) that

$$n^{-\alpha k}X_{k,n}=\sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k}C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l}[a^{k-2}(a^2+2a\sum_{q=1}^{l}m_q\frac{\eta_{n+q-1}}{n^{\varepsilon}})+\mathcal{O}(\frac{1}{n^{2\varepsilon}})], \tag{3.36}$$

which by the independence of $\{\zeta_n\}$ implies that

$$\begin{aligned}&\sigma_{k_i,k_j}(1+h,1)\\ =&4a^{k_i+k_j-2}Var(\xi)\sum_{\substack{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_{k_i}\\(l',\overrightarrow{m}'_{l'},\overrightarrow{n}'_{l'})\in\Psi_{k_j}}}C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l}C_{l'}^{\overrightarrow{m}'_{l'},\overrightarrow{n}'_{l'}}\sum_{\substack{1\le q\le l\\1\le q'\le l'}}m_qm'_{q'}\delta_{q+h,q'}.\end{aligned} \tag{3.37}$$

We shall compute explicitly the covariance $\Lambda(i,j)$ in (3.11). For convenience, we place the leftmost vertex of the circuit π' of type $(l',\overrightarrow{m}'_{l'},\overrightarrow{n}'_{l'})$ at 1. Then, the circuit π of the type $(l,\overrightarrow{m}_l,\overrightarrow{n}_l)$ contributing to (3.37) is the one with the leftmost vertex $1+h$, and $\sum_{1\le q\le l}\sum_{1\le q'\le l'}m_qm'_{q'}\delta_{q+h,q'}=$

[4] In [29, (3.5)], m_2^{k+l}, $m_2^{\mathbb{I}_h(\gamma_1)+\mathbb{I}_h(\gamma_2)}$ and the denominator $\alpha(k+l)$ shall be modified by $m_2^{\frac{1}{2}(k+l)}$, $m_2^{\frac{1}{2}(\mathbb{I}_i(\gamma_1)+\mathbb{I}_i(\gamma_2))}$ and $\alpha(k+l)+1-2\alpha$ respectively.

$1/4\sum_{q\geq 0}\mathbb{I}_{h+q}(\pi)\mathbb{I}_{h+q}(\pi')$, where $\mathbb{I}_i(\gamma)$ dentes the number of edges $\{i,i+1\}$ in the circuit γ, i.e. $\mathbb{I}_i(\gamma)=2m_{i+1-j}$ if the leftmost vertex of γ is j (see also [29, p.187]). Hence, the summation in $\sigma_{k_i,k_j}(1+h,1)$ consists of the overlapped parts of all these two circuits π and π'. Similar argument holds for $\sigma_{k_i,k_j}(1,1+h)$, with the slight modification that the leftmost vertex of π' is $1-h$ (we here extend the positions of vertices to the negative integers). Thus, using the notation $\Gamma(k,l)$ in [29, (3.1)], we deduce that the summation in (3.11) is equal to $\frac{1}{4}\sum_{(\gamma_1,\gamma_2)\in\Gamma(k_i,k_j)}\sum_{i<0}\mathbb{I}_i(\gamma_1)\mathbb{I}_i(\gamma_2)=\frac{1}{4}k_ik_j\binom{k_i}{k_i/2}\binom{k_j}{k_j/2}$ (see [29, p.213] for the equality), which together with the coefficients $(\alpha(k_i+k_j)+1-2\varepsilon)^{-1}$ and $4a^{k_i+k_j-2}Var(\xi)$ in (3.11) and (3.37) respectively yields (3.35) when k_i,k_j are even.

When k is odd, (2.2) implies that $|\overrightarrow{n}_l|$ shall be also odd. As in the preceding arguments, (3.31)-(3.33) yield that the only case contributing to the main order is that $|\overrightarrow{n}_l|=1$, namely, there is only one loop in the circuit. We use the q-th normal basis $\overrightarrow{e}_q$ of $\mathbb{R}^{[\frac{k}{2}]+1}$ to indicate that the vertex of this loop is $n+q$ if the leftmost vertex of the circuit is n, $0\leq q\leq[\frac{k}{2}]$. Then, (3.31) yields that

$$n^{-\alpha k}X_{k,n}=\sum_{\substack{(l,\overrightarrow{m}_l,\overrightarrow{e}_q)\in\Psi_k\\ 0\leq q\leq[\frac{k}{2}]}}C_l^{\overrightarrow{m}_l,\overrightarrow{e}_q}a^{k-1}\zeta_{n+q}n^{-\varepsilon}+\mathcal{O}(n^{-2\varepsilon}). \tag{3.38}$$

(Note that, this includes the easy case $k=1$.) Hence, it follows that

$$\begin{aligned}&\sigma_{k_i,k_j}(1+h,1)\\ =&a^{k_i+k_j-2}Var(\zeta)\sum_{\substack{(l,\overrightarrow{m}_l,\overrightarrow{e}_q)\in\Psi_{k_i}\\ 0\leq q\leq[\frac{k}{2}]}}\ \sum_{\substack{(l',\overrightarrow{m}'_l,\overrightarrow{e}_{q'})\in\Psi_{k_j}\\ 0\leq q'\leq[\frac{k}{2}]}}C_l^{\overrightarrow{m}_l,\overrightarrow{e}_q}C_{l'}^{\overrightarrow{m}'_{l'},\overrightarrow{e}_{q'}}\delta_{q+h,q'}.\end{aligned} \tag{3.39}$$

In particular, $\sigma_{k_i,k_j}(1+h,1)$ vanishes when $0<\varepsilon<\alpha$, since in that case $\zeta=0$.

The computation of $\Lambda(i,j)$ in this case is easier. Similarly, place the leftmost vertex of π' of the type $(l',\overrightarrow{m}'_l,\overrightarrow{n}'_l)$ at 1. (3.39) indicates that leftmost vertex of π of the type $(l,\overrightarrow{m}_l,\overrightarrow{n}_l)$ is $1+h$, and the summation in (3.39) consists of all these two circuits π and π' with exactly one loop at the same vertex. Thus, the summation in (3.11) is equal to $|\Gamma(k_i,k_j)|=k_ik_j\binom{k_i-1}{(k_i-1)/2}\binom{k_j-1}{(k_j-1)/2}$ (see [29, Remark 4] for this equality), which yields consequently (3.35).

Finally, as all ξ_n and ζ_n are independent, it follows from (3.36) and (3.38) that $\sigma_{k_i,k_j}(1+h,1) = 0$ in the other cases, which implies immediately $\Lambda(i,j) = 0$. The proof is complete. $\square$

β-Hermite ensemble ([17, 18]). This model has been widely studied in the literature. In this model, $a_i = b_i$, a_i is distributed as $\beta^{-1/2}\chi_{i\beta}$ ($\chi_{i\beta}$ is the χ distribution with $i\beta$ degrees of freedom), $\beta > 0$, d_i is normally distributed as $N(0, 2/\beta)$, and $\{a_i, d_i\}$ are all independent. One significant fact is that the density of the eigenvalue distribution is given by $C_{n\beta}\prod_{i<j}|\lambda_i - \lambda_j|^\beta\, exp(-\beta\sum_{i=1}^n \lambda_i^2/2)$ with $C_{n\beta}$ a normalization. Hence, it generalizes $\beta = 1, 2, 4$ in the classical Gaussian ensembles (i.e. GOE, GUE, GSE respectively) to continuous exponents $\beta > 0$, which are connected to lattice gas theory, Selberg-type integrals, Jack polynomials and so on. Now, using the fact that $\chi_r - \sqrt{r} \xrightarrow{d} N(0, 1/2)$, we deduce that in Assumption $(H.2)$ $\alpha = \varepsilon = 1/2$, $a = 1$, $\eta = N(0, 1/(2\beta))$, $d = 0$ and $\zeta = N(0, 2/\beta)$. Thus, Corollary (3.6) implies Gaussian fluctuations of the traces and the covariances are given by

$$\Lambda(i,j) = \begin{cases} \frac{1}{\beta}\frac{k_ik_j}{k_i+k_j}\binom{k_i}{k_i/2}\binom{k_j}{k_j/2}, & \text{if } k_i, k_j \text{ even;} \\ \frac{4}{\beta}\frac{k_ik_j}{k_i+k_j}\binom{k_i-1}{(k_i-1)/2}\binom{k_j-1}{(k_j-1)/2}, & \text{if } \varepsilon = \alpha,\ k_i, k_j \text{ odd ;} \\ 0, & \text{otherwise,} \end{cases}$$

which coincides with [29, Corollary 2][5] and [18, Theorem 1.2].

4 Deviations

This section is devoted to the large deviation and moderate deviation principles for the traces. Due to technical reasons, we will consider the tridiagonal random matrix (1.1) with entries satisfying Assumption $(H.3)$ in the case $\alpha = 0$ below.

(H.3) (i). In the non-symmetric case, $\{(a_{i-1}, d_i, b_i)\}$ are i.i.d. random vectors. Set $\nu := \mathbb{P}\circ(a_1, d_2, b_2)^{-1} \in \mathscr{P}(\mathbb{R}^3)$.

In the symmetric case ($a_i = b_i, \forall i \geq 1$), $\{a_i\}$ is a sequence of i.i.d. random variables, and $\{d_i\}$ satisfies:

[5] In [29, Corollary 2], the terms $\frac{kl}{\alpha(k+l)}$ in the even and odd cases shall be modified by $\frac{kl\sigma_Z^2}{\alpha(k+l)+1-2\varepsilon}$ and $\frac{kl\sigma_d^2}{\alpha(k+l)+1-2\alpha}$ respectively.

(*ii*). $\{d_i\}$ is a sequence of i.i.d. random variables and independent of $\{a_i\}$. Set $\nu' := \mathbb{P} \circ (d_1, a_1)^{-1} \in \mathscr{P}(\mathbb{R}^2)$. Or,

(*iii*). $d_i = f(a_{i-1}, a_i)$, where f is a continuous function and $a_0 = 0$. Set $\nu'' = \mathbb{P} \circ (a_1)^{-1} \in \mathscr{P}(\mathbb{R})$.

Moreover, in all cases $\{a_i, d_i, b_i\}$ are bounded random variables.

Remark 4.1 *Assumption* (*H*.3) *already includes the models corresponding to the case* $\alpha = 0$ *in Subsection 3.4, such as the Anderson model, the Hatano-Nelson model, the random birth-death Markov kernel and the random birth-death* Q *matrix.*

4.1 Large deviations

Let us first introduce some basic notations and definitions for large deviation principles. More details can be found in [13] and [20]. Then, we formulate and prove our result in Theorem 4.2 below.

Consider a complete separable metric space $\mathscr{X}$, $\mathscr{P}(\mathscr{X})$ denotes all the probability measures in $\mathscr{X}$. A sequence of probability measures $\{\mu_n\} \subset \mathscr{P}(\mathscr{X})$ is said to satisfy the large deviation principle (LDP for short) with speed $s_n \to \infty$ and good rate function $I : \mathscr{X} \to [0, \infty]$, if the level sets $\{x \in \mathscr{X} : I(x) \le c\}$ are compact for all $c \in [0, \infty)$ and if for all Borel set A of $\mathscr{X}$,

$$-\inf_{x \in A^o} I(x) \le \liminf_{n\to\infty} \frac{1}{s_n} \log \mu_n(A) \le \limsup_{n\to\infty} \frac{1}{s_n} \log \mu_n(A) \le -\inf_{x \in \overline{A}} I(x),$$

where A^o and $\overline{A}$ denote the interior and closure of A respectively. In that case, we shall simply say that $\{\mu_n\}$ satisfies the $LDP(s_n, I)$ on $\mathscr{X}$. We also say that a family of $\mathscr{X}$-valued random variables satisfies the $LDP(s_n, I)$ if the family of their laws does. For any $\mu, \mu' \in \mathscr{P}(\mathscr{X})$, the relative entropy of μ with respect to μ' is defined by

$$H(\mu|\mu') = \begin{cases} \int g \log g d\mu', & \text{if } g = \frac{d\mu}{d\mu'} \text{ exists;} \\ \infty, & \text{otherwise.} \end{cases}$$

We also consider the product space $\mathscr{X}^r$ with $r \ge 2$. A probability measure $\mu \in \mathscr{P}(\mathscr{X}^r)$ is said to be shift invariant if for any Borel set A of $\mathscr{X}^{r-1}$,

$$\mu(x \in \mathscr{X}^r : (x_1, \cdots, x_{r-1}) \in A) = \mu(x \in \mathscr{X}^r : (x_2, \cdots, x_r) \in A).$$

Moreover, for any $\mu \in \mathscr{P}(\mathscr{X}^{r-1})$, $\rho \in \mathscr{P}(\mathscr{X})$, define the probability measure $\mu \otimes_r \rho \in \mathscr{P}(\mathscr{X}^r)$ by, for any Borel set A of $\mathscr{X}^r$,

$$(\mu \otimes_r \rho)\ (A) = \int\limits_{\mathscr{X}^{r-1}} \mu(dx) \int\limits_{\mathscr{X}} I_{\{(x,y)\in A\}}\rho(dy),$$

where $I_{\{(x,y)\in A\}}$ is the indicator function of the set $\{(x, y) \in A\}$. Define for $\rho \in \mathscr{P}(\mathscr{X})$,

$$I_{r,\rho}(\mu) = \begin{cases} H(\mu|\mu_{r-1} \otimes_r \rho), & \text{if } \mu \text{ is shift invariant;} \\ \infty, & \text{otherwise,} \end{cases} \tag{4.1}$$

where μ_{r-1} denotes the marginal of μ on the first $(r-1)$ coordinates.

Now, come to the tridiagonal random matrix (1.1). For each $k \geq 1$, set $r_k = [\frac{k}{2}] + 1$. In the non-symmetric case, take $\mathscr{X} = \mathbb{R}^3$, and define the function F on $(\mathbb{R}^3)^{r_k}$ by

$$F(\alpha_1, \cdots, \alpha_{r_k}) = \sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k} C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l} \prod_{j=0}^{l-1}(x_{1+j}z_{1+j})^{m_{j+1}} \prod_{j=0}^{l} y_{1+j}^{n_j}, \tag{4.2}$$

where $\alpha_i = (x_{i-1}, y_i, z_i) \in \mathbb{R}^3$, $1 \leq i \leq r_k$ and $x_0 = 0$. Associating with F, define $I^F_{3r_k,\nu}$ on $\mathbb{R}^3$ by

$$I^F_{3r_k,\nu}(x) := \inf\{I_{r_k,\nu}(\mu) :\ \mu \in \mathscr{P}((\mathbb{R}^3)^{r_k}),\ x = \mu(F)\},$$

where ν is defined as in Assumption $(H.3)$, and $\mu(F)$ denotes the integration of F with respect to μ.

Similarly, in the symmetric case (ii) in $(H.3)$, take $\mathscr{X} = \mathbb{R}^2$, define the function F' on $(\mathbb{R}^2)^{r_k}$ by

$$F'(\beta_1, \cdots, \beta_{r_k}) = \sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k} C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l} \prod_{j=0}^{l-1} z_{1+j}^{2m_{j+1}} \prod_{j=0}^{l} y_{1+j}^{n_j} \tag{4.3}$$

with $\beta_i = (y_i, z_i)$, $1 \leq i \leq r_k$, and the corresponding function $I^{F'}_{2r_k,\nu'}$ is defined by

$$I^{F'}_{2r_k,\nu'}(x) := \inf\{I_{r_k,\nu'}(\mu) :\ \mu \in \mathscr{P}((\mathbb{R}^2)^{r_k}),\ x = \mu(F')\}$$

where ν' is defined as in Assumption $(H.3)$.

In the symmetric case (iii) in $(H.3)$, take $\mathscr{X} = \mathbb{R}$, and define the function F'' on $\mathbb{R}^{r_k+1}$ by

$$F''(z_0, \cdots, z_{r_k}) = F'(\gamma_1, \cdots, \gamma_{r_k}),$$

where $\gamma_i = (f(z_{i-1}, z_i), z_i)$, $1 \leq i \leq r_k$. The related function $I^{F''}_{r_k+1,\nu''}$ is defined analogously by

$$I^{F''}_{r_k+1,\nu''}(x) := \inf\{I_{r_k+1,\nu''}(\mu) : \ \mu \in \mathscr{P}((\mathbb{R})^{r_k+1}), \ x = \mu(F'')\}$$

with ν'' defined as in (iii) in $(H.3)$.

We are ready to state our large deviation result. Since when $k = 1$, $TrQ_n^k = \sum_{i=1}^n d_i$, i.e. the sum of i.i.d random variables, of which the large deviation is well known by the Cramér theorem (see [13, Theorem 2.2.3]). Therefore, below we are concerned with the case $k \geq 2$.

Theorem 4.2 *Assume $(H.3)$. Let $k \geq 2$ and set $r_k := [\frac{k}{2}] + 1$. Then, $\{\frac{1}{n}TrQ_n^k\}_{n\geq 1}$ satisfies the $LDP(n, I^F_{3r_k,\nu})$, $LDP(n, I^{F'}_{2r_k,\nu'})$ and $LDP(n, I^{F''}_{r_k+1,\nu''})$ in the cases $(i) - (iii)$ in $(H.3)$ respectively.*

Proof. Let us first consider the non-symmetric case (i) in $(H.3)$. We note that $\{\frac{1}{n}TrQ_n^k\}$ and $\{\frac{1}{n}\sum_{i=1}^n X_{k,i}\}$ is exponentially equivalent, i.e., for any $\delta > 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\frac{1}{n}|TrQ_n^k - \sum_{i=1}^n X_{k,i}| > \delta) = -\infty. \tag{4.4}$$

(See the Appendix for the proof.) Thus, by the exponential equivalence theorem([13, Theorem 4.2.13]), the proof reduces to proving the large deviation of $\frac{1}{n}\sum_{i=1}^n X_{k,i}$.

Now, set the random vectors $\alpha_i = (a_{i-1}, d_i, b_i) \in \mathbb{R}^3$, $i \geq 1$. By the independence and the identical distribution of $\{\alpha_i\}$ in Assumption $(H.3)$, we regard $\{\alpha_i\}$ as a Markov chain on $\mathbb{R}^3$ with the transition probability $\pi(x, dy) := \nu(dy)$, $x, y \in \mathbb{R}^3$. Note that $\{\pi(x, dy)\}$ obviously satisfy the uniform Assumption (U) in [13, p.275] (see also Hypothesis 1.1(a) in [20]), i.e., $\{\alpha_i\}$ is a uniform Markov chain on $\mathbb{R}^3$. Thus, applying [13, Theorem 6.5.12] (see also [20, Theorem 1.4]), we have that the multivariate empirical measure $\mu_n := \frac{1}{n}\sum_{i=1}^n \delta_{(\alpha_i,\cdots,\alpha_{i+r_k-1})}$ satisfies the $LDP(n, I_{r_k,\nu})$ with the good rate function $I_{r_k,\nu}$ defined as in (4.1).

On the other hand, by the definition of $X_{k,i}$ in (3.12), we note that

$$X_{k,i} = F(\alpha_i, \cdots, \alpha_{i+r_k-1}), \tag{4.5}$$

namely, $X_{k,i}$ can be viewed as a continuous function on the product space $(\mathbb{R}^3)^{r_k}$. Moreover,

$$\frac{1}{n}\sum_{i=1}^{n} X_{k,i} = \frac{1}{n}\sum_{i=1}^{n} F(\alpha_i, \cdots, \alpha_{i+r_k-1}) = \mu_n(F),$$

which implies that $\frac{1}{n}\sum_{i=1}^{n} X_{k,i}$ is an additive functional of the uniform Markov chain.

Therefore, we can apply the contraction principle (see [13, Theorem 4.2.1]) to obtain the large deviation of $\frac{1}{n}\sum_{i=1}^{n} X_{k,i}$ as specified in the non-symmetric case.

The symmetric case can be treated analogously. In fact, in the case (ii) in $(H.3)$, $X_{k,i} = F'(\beta_i, \cdots, \beta_{i+r_k-1})$, and the random vectors $\beta_i = (d_i, b_i) \in \mathbb{R}^2$ forms a uniform Markov chain on $\mathbb{R}^2$ with the transition probability $\pi'(x, dy) = \nu'(dy)$. Moreover, in the case (iii) in $(H.3)$, $X_{k,i} = F''(b_{i-1}, \cdots, b_{i+r_k-1})$, and $\{b_i\} \subset \mathbb{R}$ forms a uniform Markov chain in $\mathbb{R}$ with $\pi''(x, dy) = \nu''(dy)$. Therefore, applying [13, Theorem 6.5.12] and the contraction principle, we obtain the asserted large deviation results. The proof of Theorem 4.2 is complete. $\square$

4.2 Moderate deviations

Moderate deviation principles for dependent random variables are widely studied in the literature, see e.g. [32, 11, 27] and references therein. Here, for the moderate deviations of the traces, we prefer to give an elementary proof based on the blocking arguments as those in the proof of Theorem 3.2.

Theorem 4.3 *Assume $(H.3)$. Set $\lambda_n = n^{-\nu}$ with $\nu \in (0,1)$, and let $\widetilde{TrQ_n^k} = TrQ_n^k - \mathbb{E}TrQ_n^k$. Then, for every $k \geq 1$ and $D_k > 0$, where D_k is defined as in (3.34), $\{\sqrt{\frac{\lambda_n}{n}}\widetilde{TrQ_n^k}\}_{n\geq 1}$ satisfies the $LDP(\lambda_n^{-1}, \frac{x^2}{2D_k})$.*

Proof of Theorem 4.3. Let $\widetilde{X}_{k,i} = X_{k,i} - \mathbb{E}X_{k,i}$. We first note that $\{\sqrt{\frac{\lambda_n}{n}}\widetilde{TrQ_n^k}\}$ and $\{\sqrt{\frac{\lambda_n}{n}}\sum_{i=1}^{n}\widetilde{X}_{k,i}\}$ is exponentially equivalent, i.e. for any

$\delta > 0$,

$$\limsup_{n\to\infty} \lambda_n \log \mathbb{P}\left(\sqrt{\frac{\lambda_n}{n}}|Tr\widetilde{Q_n^k} - \sum_{i=1}^n \widetilde{X}_{k,i}| \geq \delta\right) = -\infty. \tag{4.6}$$

(The proof is postponed to the Appendix.) Hence, it is equivalent to prove the moderate deviation for $\sum_{i=1}^n \widetilde{X}_{k,i}$. For this purpose, we will use the blocking arguments as in the proof of Theorem 3.2.

Let $p = n^{\nu+\varepsilon}$, where $\varepsilon \in (\frac{3}{4}(1-\nu), 1-\nu)$. Set $n' = [\frac{n}{p}]$ and $r = n - pn'$. Let $\widetilde{Y}_{n,i}, \widetilde{U}_{n,i}, \widetilde{Z}_{n,i}$ and $\widetilde{T}_n$ be as in the proof of Theorem 3.2, but with $\alpha = 0$. Then

$$\sqrt{\frac{\lambda_n}{n}} \sum_{i=1}^n \widetilde{X}_{k,i} = \sqrt{\frac{\lambda_n}{n}} \sum_{i=1}^p \widetilde{U}_{n,i} + \sqrt{\frac{\lambda_n}{n}} \widetilde{T}_n.$$

Moreover, denote by Λ_n the logarithmic moment generating function of $\sqrt{\frac{\lambda_n}{n}} \sum\limits_{i=1}^p \widetilde{U}_{n,i}$, i.e. $\Lambda_n(t) = \log \mathbb{E} \exp(t\sqrt{\frac{\lambda_n}{n}} \sum_{i=1}^p \widetilde{U}_{n,i})$, $t \in \mathbb{R}$.

We shall prove below that, for every $\delta > 0$,

$$\limsup_{n\to\infty} \lambda_n \log \mathbb{P}\left(|\sqrt{\frac{\lambda_n}{n}}\widetilde{T}_n| \geq \delta\right) = -\infty, \tag{4.7}$$

and

$$\lim_{n\to\infty} \lambda_n \Lambda_n(\lambda_n^{-1} t) = \frac{t^2}{2} D_k, \tag{4.8}$$

where D_k is the variance defined as in (3.34). Then, by the exponential equivalence, (4.7) implies that we only need to consider the moderate deviations of $\sum_{i=1}^p \widetilde{U}_{n,i}$. Consequently, (4.8) and the Gärtner-Ellis theorem (see e.g. [13]) yield the asserted moderate deviation principle for the traces.

It remains to prove (4.7) and (4.8). For the proof of (4.7), since $\widetilde{T}_n = \sum_{i=1}^p \widetilde{Z}_{n,i}$ and $\widetilde{Z}_{n,i}$ are independent, using Assumption $(H.3)$ and the Bernstein inequality (see e.g. [2, p.21]), we have

$$\mathbb{P}(|\sqrt{\frac{\lambda_n}{n}}\widetilde{T}_n| \geq \delta) \leq 2e^{-\delta^2/(2(B_n^2 + c\delta))},$$

where $B_n^2 = \frac{\lambda_n}{n}\mathbb{E}|\widetilde{T}_n|^2 = \mathcal{O}(\frac{\lambda_n p}{n})$ and $c = \sqrt{\frac{\lambda_n}{n}} \sup_{1\le i\le p} \|\widetilde{Z}_{n,i}\|_\infty = \mathcal{O}(\sqrt{\frac{\lambda_n}{n}})$. Thus,

$$\begin{aligned}\lambda_n \log \mathbb{P}(|\sqrt{\frac{\lambda_n}{n}}\widetilde{T}_n| \ge \delta) \le &\lambda_n \log(2e^{-\delta^2/(2(B_n^2+c\delta))}) \\ =&\lambda_n \log 2 - \frac{\delta^2}{\mathcal{O}(\frac{p}{n} + \frac{1}{\sqrt{\lambda_n n}})} \to -\infty,\end{aligned}$$

which implies (4.7), as claimed.

Coming to the proof of (4.8), we note that by the uniform boundedness in Assumption $(H.3)$ and $\varepsilon > \frac{1}{2}(1-\nu)$,

$$\frac{1}{\sqrt{\lambda_n n}}\widetilde{U}_{n,i} = \mathcal{O}(n'/\sqrt{\lambda_n n}) = \mathcal{O}(n^{\frac{1}{2}(1-\nu)-\varepsilon}) = o(1).$$

Hence, it follows that

$$\begin{aligned}\lambda_n \Lambda_n(\lambda_n^{-1}t) =& \sum_{i=1}^p \lambda_n \log \mathbb{E}\exp(\frac{t}{\sqrt{\lambda_n n}}\widetilde{U}_{n,i}) \\ =& \sum_{i=1}^p \lambda_n \log\left[1 + \frac{t^2}{2\lambda_n n}\mathbb{E}\widetilde{U}_{n,i}^2 + \mathcal{O}(\frac{\mathbb{E}\widetilde{U}_{n,i}^3}{(\lambda_n n)^{\frac{3}{2}}})\right]. \end{aligned} \tag{4.9}$$

Similarly to (3.18) and (3.19), we have

$$\begin{aligned}\frac{1}{\lambda_n n}\mathbb{E}\widetilde{U}_{n,i}^2 =& \frac{1}{\lambda_n n}\left[\mathbb{E}\left(\sum_{h=1}^{m_k}\widetilde{Y}_{n,(i-1)n'+h}\right)^2 + \sum_{h=1}^{n'-2m_k} A_{n,(i-1)n'+h}\right] \\ =& \mathcal{O}(\frac{1}{\lambda_n n}) + \frac{1}{\lambda_n n}\sum_{h=1}^{n'-2m_k} A_{n,(i-1)n'+h},\end{aligned}$$

where $A_{n,(i-1)n'+h}$ are defined as in (3.17). Hence, it follows from Assumption $(H.3)$ and similar computations as in (3.19)-(3.23) that

$$\frac{1}{\lambda_n n}\mathbb{E}\widetilde{U}_{n,i}^2 = \mathcal{O}(\frac{1}{\lambda_n n}) + \frac{1}{\lambda_n p}(D_k + o(1)). \tag{4.10}$$

Moreover, for the third moment in (4.9), by the choices of λ_n and n',

$$\begin{aligned}\frac{1}{(\lambda_n n)^{\frac{3}{2}}}\mathbb{E}\widetilde{U}_{n,i}^3 =&\frac{1}{(\lambda_n n)^{\frac{3}{2}}}\mathbb{E}(\sum_{j=(i-1)n'+1}^{in'-m_k}\widetilde{Y}_{n,j})^3\\ =&\frac{1}{(\lambda_n n)^{\frac{3}{2}}}\mathcal{O}((n')^3)\\ =&n^{3(\frac{1}{2}-\frac{1}{2}\nu-\varepsilon)}=o(1). \end{aligned}\tag{4.11}$$

Consequently, plugging (4.10) and (4.11) into (4.9), we obtain that

$$\begin{aligned}\lambda_n\Lambda_n(\lambda_n^{-1}t)=&\sum_{i=1}^{p}\lambda_n\left[\frac{t^2}{2\lambda_n n}\mathbb{E}\widetilde{U}_{n,i}^2+\mathcal{O}(\frac{\mathbb{E}\widetilde{U}_{n,i}^3}{(\lambda_n n)^{\frac{3}{2}}})\right]+o(1)\\ =&\mathcal{O}(\frac{1}{n'})+\frac{t^2}{2}D_k+\mathcal{O}(n^{\frac{3}{2}-\frac{3}{2}\nu-2\varepsilon})+o(1)\\ \to&\frac{t^2}{2}D_k.\end{aligned}$$

which implies (4.8), thereby completing the proof. □

5 Discussions

1. For more general finite diagonal random matrix (see e.g. [6], [29]), the Gaussian fluctuations and deviations of the traces can be treated in a similar way. Indeed, using the finite width of band, we can still reduce the asymptotical analysis of the traces to those of the corresponding m-dependent random variables.

2. It seems difficult to prove deviation results for the general case $\alpha > 0$. In some special cases, e.g. the GUE case ($\alpha = 1/2$), deviation results are known for the empirical spectral distribution (hence for the traces by the contraction principle), we refer to [1, 12].

3. In the derivation of the large deviation results in Subsection 4.1, we regard the trace as an additive functional of a uniform Markov chain. With this point of view, one can expect to achieve similar results for the traces when the entries of (1.1) form an appropriate Markov chain, e.g. positive Harris recurrent. We refer the interested reader to [10] and references therein.

6 Appendix

Proof of (3.13). By Assumption $(H.1)$ and $(H.2)$,

$$\sup_{i\geq 1}\mathbb{E}|n^{-\alpha k}\widetilde{Q}_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l}|^2<\infty, \tag{6.1}$$

where $\widetilde{Q}_{l,i}^{\overrightarrow{m}_l}=Q_{l,i}^{\overrightarrow{m}_l}-\mathbb{E}Q_{l,i}^{\overrightarrow{m}_l}$. Hence,

$$\begin{aligned}
&\|n^{-\alpha k}(Tr\widetilde{Q_n^k}-\sum_{i=1}^n\widetilde{X}_{k,i})\|_2\\
=&\|\sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k}C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l}\sum_{i=n-l+1}^n n^{-\alpha k}\widetilde{Q}_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l}\|_2\\
\leq&\sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k}C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l}\sum_{i=n-l+1}^n\|n^{-\alpha k}\widetilde{Q}_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l}\|_2=\mathcal{O}(1),
\end{aligned}$$

with $\|\cdot\|_2$ denoting the standard L^2 norm, which implies (3.13). □

Proof of (3.21). It is equivalent to prove that, as $n\to\infty$,

$$n^{-2\alpha k}Cov(X_{k,n},X_{k,n+j})\to Cov(Z_{k,1},Z_{k,1+j}). \tag{6.2}$$

First consider the nonsymmetic case. Let $\alpha_i=(a_{i-1},d_i,b_i)$, $i\geq 1$ and $a_0=0$. By the independence and weak convergence of α_i in Assumption $(H.1)$, it is not difficult to deduce that

$$n^{-\alpha k}(\alpha_n,\cdots,\alpha_{n+[\frac{k}{2}]})\xrightarrow{d}(\widetilde{\alpha}_1,\cdots,\widetilde{\alpha}_{1+[\frac{k}{2}]}),$$

where $\widetilde{\alpha}_i$ are independent but with the same distribution as that of $(a+\eta,d+\zeta,b+\xi)$. Then, by (4.5) and the continuous mapping theorem ([19, Theorem 3.2.4]),

$$n^{-\alpha k}X_{k,n}\xrightarrow{d}F(\widetilde{\alpha}_1,\cdots,\widetilde{\alpha}_{1+[\frac{k}{2}]}) \tag{6.3}$$

with F the continuous function defined as in (4.2).

Similarly,

$$n^{-\alpha k}X_{k,n+j}\xrightarrow{d}F(\widetilde{\alpha}_{1+j},\cdots,\widetilde{\alpha}_{1+j+[\frac{k}{2}]}). \tag{6.4}$$

On the other hand, by (3.3) and Hölder's inequality

$$\sup_{n\geq 1}\mathbb{E}(n^{-\alpha k}X_{k,n})^4<\infty, \tag{6.5}$$

which implies the uniform integrabilities of $n^{-2\alpha k}X_{k,n}X_{k,n+j}$, $n^{-\alpha k}X_{k,n}$ and $n^{-\alpha k}X_{k,n+j}$, $n\geq 1$.

Therefore, by (6.3)-(6.5), we can apply the Skorohod representation theorem and the uniform integrability to take the limit and obtain (6.2) for the non-symmetric case. (See e.g. [19, Theorem 3.2.4] for similar arguments for the bounded continuous mapping.)

The symmetric case can be proved analogously. In fact, with the uniform integrability (6.5), we only need to check the weak convergence of $n^{-\alpha k}X_{k,n}$.

In the case that all d_i are independent of $a_i(=b_i)$, let $\beta_i=(d_i,a_i)$, $i\geq 1$. In this case, $X_{k,n}=F'(\beta_n,\cdots,\beta_{n+[\frac{k}{2}]})$ with F' defines as in (4.3). Then, it follows from similar arguments as above that

$$n^{-\alpha k}X_{k,n}\xrightarrow{d}F'(\widetilde{\beta}_1,\cdots,\widetilde{\beta}_{1+[\frac{k}{2}]}), \tag{6.6}$$

where $\widetilde{\beta}_i$, $1\leq i\leq[\frac{k}{2}]+1$, are independent but with the common distribution as that of $(d+\xi,a+\eta)$.

In the case that $d_i=f(a_{i-1},a_i)$, then $\beta_i=(f(a_{i-1},a_i),a_i)$, $X_{k,n}$ is now a continuous function of $a_{n-1},\cdots,a_{n+[\frac{k}{2}]}$. As $(a_{n-1},\cdots,a_{n+[\frac{k}{2}]})\xrightarrow{d}(\widetilde{\alpha_1},\cdots,\widetilde{\alpha}_{[\frac{k}{2}]+2})$, by the continuous mapping theorem, we can obtain the weak convergence of $n^{-\alpha k}X_{k,n}$. The proof is consequently complete. □

Proof of (4.4). Set $N(k)=|\Psi_k|$, the number of sets in Ψ_k which is defined as in (2.3). $N(k)$ is finite and depends only on k. Note that

$$\begin{aligned}&\mathbb{P}\left(\frac{1}{n}\left|TrQ_n^k-\sum_{i=1}^n X_i\right|\geq\delta\right)\\ \leq&\sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k}\mathbb{P}\left(\left|C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l}\sum_{i=n-l+1}^{n}Q_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l}\right|\geq\frac{n\delta}{N(k)}\right).\end{aligned}$$

Then, letting C denote the maximum of $C_l^{\overrightarrow{m}_l,\overrightarrow{n}_l}$ over the finite sets in Ψ_k and

setting $\delta_k = \delta/([\frac{k}{2}]CN(k))$, we deduce that

$$\mathbb{P}\left(\frac{1}{n}\left|TrQ_n^k - \sum_{i=1}^{n} X_{k,i}\right| \geq \delta\right) \leq \sum_{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k} \sum_{i=n-l+1}^{n} \mathbb{P}\left(\left|Q_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l}\right| \geq n\delta_k\right).$$

Since by (2.1) and Assumption $(H.3)$, $\{|Q_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l}|\}_{i\geq 1}$ is uniformly bounded, which implies that $\mathbb{P}(|Q_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l}| \geq n\delta_k) = 0$ for n large enough. Hence, by Lemma 1.2.15 in [13],

$$\begin{aligned}
&\frac{1}{n}\log\mathbb{P}\left(\frac{1}{n}|TrQ_n^k - \sum_{i=1}^{n} X_{k,i}| \geq \delta\right) \\
&\leq \max_{\substack{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k \\ 1\leq j\leq [\frac{k}{2}]}} \frac{1}{n}\log\mathbb{P}\left(|Q_{l,n-l+j}^{\overrightarrow{m}_l,\overrightarrow{n}_l}| \geq n\delta_k\right) = -\infty. \qquad (6.7)
\end{aligned}$$

yielding (4.4) as claimed. □

Proof of (4.6). Similarly to the proof of (6.7), we derive that

$$\begin{aligned}
&\lambda_n \log\mathbb{P}\left(\sqrt{\frac{\lambda_n}{n}}|Tr\widetilde{Q_n^k} - \sum_{i=1}^{n} \widetilde{X}_{k,i}| \geq \delta\right) \\
&\leq \max_{\substack{(l,\overrightarrow{m}_l,\overrightarrow{n}_l)\in\Psi_k \\ 1\leq j\leq [\frac{k}{2}]}} \lambda_n \log\mathbb{P}\left(|\widetilde{Q}_{l,n-l+j}^{\overrightarrow{m}_l,\overrightarrow{n}_l}| \geq \sqrt{\frac{n}{\lambda_n}}\delta_k\right)
\end{aligned}$$

with δ_k defined as in the proof of (6.7), which yields (4.6), due to the fact that $\{Q_{l,i}^{\overrightarrow{m}_l,\overrightarrow{n}_l}\}_{i\geq 1}$ is uniformly bounded and $n/\lambda_n \to \infty$.

Acknowledgement. The author is grateful to Professor Dong Han for several valuable discussions.

References

[1] G. Anderson, A. Guionnet, O. Zeitouni, An introduction to random matrices. Cambridge Studies in Advanced Mathematics, 118. Cambridge University Press, Cambridge, 2010.

[2] Z. D. Bai, J. Silverstein, Spectral analysis of large dimensional random matrices. Second edition. Springer Series in Statistics. Springer, New York, 2010.

[3] G. Ben Arous, A. Guionnet, Large deviations for Wigner's law and Voiculescu's non-commutative entropy. Probab. Theory Related Fields 108 (1997), no. 4, 517-542.

[4] G. Ben Arous, O. Zeitouni, Large deviations from the circular law. ESAIM Probab. Statist. 2 (1998), 123-134.

[5] C. Bordenave, P. Caputo, D. Chafaï, Spectrum of large random reversible Markov chains: two examples. ALEA Lat. Am. J. Probab. Math. Stat. 7 (2010), 41-64.

[6] A. Bose, S. Sen, Finite diagonal random matrices. J. Theoret. Probab. 26 (2013), no. 3, 819-835.

[7] R. Carmona, J. Lacroix, Spectral Theory of Random Schrödinger Operators, Birkhäuser, Boston, 1990.

[8] D. Chafaï, Aspects of large random Markov kernels. Stochastics 81 (2009), no. 3-4, 415-429.

[9] K. L. Chung, A course in probability theory. Second edition. Probability and Mathematical Statistics, Vol. 21. Academic Press, New York-London, 1974.

[10] A. de Acosta, P. Ney, Large deviations for additive functionals of Markov chains. Mem. Amer. Math. Soc. 228 (2014).

[11] J. Dedecker, F. Merlevède, M. Peligrad, S. Utev, Moderate deviations for stationary sequences of bounded random variables. Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009), no. 2, 453-476.

[12] A. Dembo, A. Guionnet, O. Zeitouni, Moderate deviations for the spectral measure of certain random matrices. Ann. Inst. H. Poincaré Probab. Statist. 39 (2003), no. 6, 1013-1042.

[13] A. Dembo; O. Zeitouni, Large deviations techniques and applications. Corrected reprint of the second (1998) edition. Stochastic Modelling and Applied Probability, 38. Springer-Verlag, Berlin, 2010.

[14] P. Diaconis, M. Shahshahani, On the eigenvalues of random matrices. Studies in applied probability. J. Appl. Probab. 31A (1994), 49-62.

[15] H. Döring, P. Eichelsbacher, Moderate deviations via cumulants. J. Theoret. Probab. 26 (2013), no. 2, 360-385.

[16] H. Döring, P. Eichelsbacher, Moderate deviations for the eigenvalue counting function of Wigner matrices. ALEA Lat. Am. J. Probab. Math. Stat. 10 (2013), no. 1, 27-44.

[17] I. Dumitriu, A. Edelman, Matrix models for beta ensembles. J. Math. Phys. 43 (2002), no. 11, 5830-5847.

[18] I. Dumitriu, A. Edelman, Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models. J. Math. Phys. 47 (2006), no. 6, 063302, 36 pp.

[19] R. Durrett, Probability: theory and examples. Fourth edition. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010.

[20] R. S. Ellis, Large deviations for the empirical measure of a Markov chain with an application to the multivariate empirical measure. Ann. Probab. 16 (1988), no. 4, 1496-1508.

[21] I. Y. Goldsheid, B. A. Khoruzhenko, The Thouless formula for random non-Hermitian Jacobi matrices. Probability in mathematics. Israel J. Math. 148 (2005), 331-346.

[22] A. Guionnet, Large random matrices: lectures on macroscopic asymptotics. Lecture Notes in Mathematics, 1957. Springer, Berlin 2009.

[23] D. Han, D. Zhang, Limiting spectral distribution of random birth-death Q matrices (in Chinese). Sci. Sin. Math. 45 (2015), 539-558, doi: 10.1360/N012015-00065.

[24] W. Hoeffding, H. Robbins, The central limit theorem for dependent random variables. Duke Math. J. 15, (1948). 773-780.

[25] K. Johansson, On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. 91 (1998), no. 1, 151-204.

[26] A. Lytova, L. Pastur, Central limit theorem for linear eigenvalue statistics of random matrices with independent entries. Ann. Probab. 37 (2009), no. 5, 1778-1840.

[27] F. Merlevède, M. Peligrad, E. Rio, A Bernstein type inequality and moderate deviations for weakly dependent sequences. Probab. Theory Related Fields 151 (2011), no. 3-4, 435-474.

[28] L. A. Pastur, A. L. Figotin, Spectra of Random and Almost-periodic Operators. Springer, Berlin-Heidelberg-New York, 1992.

[29] I. Popescu, General tridiagonal random matrix models, limiting distributions and fluctuations. Probab. Theory Related Fields 144 (2009), no. 1-2, 179-220.

[30] A. Soshnikov, Gaussian fluctuation for the number of particles in Airy, Bessel, sine, and other determinantal random point fields. J. Statist. Phys. 100 (2000), no. 3-4, 491-522.

[31] A. Soshnikov, Gaussian limit for determinantal random point fields. Ann. Probab. 30 (2002), no. 1, 171-187.

[32] L. M. Wu, Moderate deviations of dependent random variables related to CLT. Ann. Probab. 23 (1995), no. 1, 420-445.